

RESEARCH ARTICLE

Open Access



# Diagnostic performance of automated, streamlined, daily updated exome analysis in patients with neurodevelopmental delay

Go Hun Seo<sup>1</sup>, Hane Lee<sup>1</sup>, Jungsul Lee<sup>1</sup>, Heonjong Han<sup>1</sup>, You Kyung Cho<sup>1</sup>, Minji Kim<sup>2</sup>, Yunha Choi<sup>2</sup>, Jeongmin Choi<sup>3</sup>, In Hee Choi<sup>3</sup>, Seonkyeong Rhie<sup>4</sup>, Kyu Young Chae<sup>4</sup>, Yoo-Mi Kim<sup>5</sup>, Chong Kun Cheon<sup>6</sup>, Su Jin Kim<sup>7</sup>, Jieun Lee<sup>7</sup>, Eungu Kang<sup>8</sup>, Jung Hye Byeon<sup>8</sup>, Hee Joon Yu<sup>9</sup>, Young-Lim Shin<sup>10</sup>, Arum Oh<sup>11</sup>, Woo Jin Kim<sup>12</sup>, Mi-Sun Yum<sup>2\*†</sup>, Beom Hee Lee<sup>2,3\*†</sup> and Baik-Lin Eun<sup>8\*†</sup>

## Abstract

**Background:** The diagnostic yield of whole-exome sequencing (WES) varies from 30%–50% among patients with mild to severe neurodevelopmental delay (NDD)/intellectual disability (ID). Routine retrospective reanalysis of undiagnosed patients has increased the total diagnostic yield by 10–15%. Here, we performed proband-only WES of 1065 patients with NDD/ID and applied a prospective, daily reanalysis automated pipeline to patients without clinically significant variants to facilitate diagnoses.

**Methods:** The study included 1065 consecutive patients from 1056 nonconsanguineous unrelated families from 10 multimedical centers in South Korea between April 2018 and August 2021. WES data were analyzed daily using automatically updated databases with variant classification and symptom similarity scoring systems.

**Results:** At the initial analysis, 402 patients from 1056 unrelated families (38.0%, 402/1,056 families) had a positive genetic diagnosis. Daily prospective, automated reanalysis resulted in the identification of 34 additional diagnostic variants in 31 patients (3%), which increased our molecular diagnostic yield to 41% (433/1056 families). Among these 31 patients, 26 were diagnosed with 23 different diseases that were newly discovered after 2019. The time interval between the first analysis and the molecular diagnosis by reanalysis was  $1.2 \pm 0.9$  years, which was shorter in the patients enrolled during the latter part of the study period.

**Conclusion:** Daily updated databases and reanalysis systems enhance the diagnostic performance in patients with NDD/ID, contributing to the rapid diagnosis of undiagnosed patients by applying the latest molecular genetic information.

**Keywords:** Whole exome sequencing, Neurodevelopmental delay, Reanalysis

\*Correspondence: misun.yum@gmail.com; bhlee@amc.seoul.kr; bleun@korea.ac.kr

†Mi-Sun Yum, Beom Hee Lee and Baik-Lin Eun contributed equally to this manuscript

<sup>2</sup> Department of Pediatrics, Asan Medical Center Children's Hospital, University of Ulsan College of Medicine, 88, Olympic-ro 43-Gil, Songpa-Gu, Seoul 05505, South Korea

<sup>3</sup> Department of Pediatrics, Korea University Guro Hospital, 148 Gurodong-ro, Guro-gu, Seoul 08308, South Korea

Full list of author information is available at the end of the article

## Introduction

The prevalence of neurodevelopmental delay (NDD)/intellectual disability (ID) is estimated to be from 3% to 17% of children aged 3 to 17 years (Ropers 2010; Vasudevan and Suri 2017; Zablotsky et al. 2019). NDD/ID are phenotypically heterogeneous among patients in terms of the range and severity and can occur alone or in combination with multiple congenital anomalies,



dysmorphic features, behavioral problems and additional neurological features such as epilepsy (Casanova et al. 2018; Lecoquierre et al. 2019). NDD/ID are also genetically heterogeneous, with more than 700 associated genes, making it difficult to pinpoint the causal gene; thus, many patients remain undiagnosed (Vissers et al. 2016; Study 2017).

Currently, chromosomal microarray (CMA) is recommended as the first-line diagnostic test for patients with NDD/ID to identify copy number variants such as deletions and duplications (Moeschler and Shevell 2014; Srour and Shevell 2014). However, the diagnostic yield of CMA is approximately 10–15%, and many NDD/ID patients remain undiagnosed (Clark et al. 2018). A significant reduction in the cost of next-generation sequencing (NGS) has led to widespread use of trio- or proband-based whole exome or genome sequencing for identifying causal variants among NDD/ID patients, and the diagnostic yield is higher compared to that of CMA (Vissers et al. 2016; Ligt et al. 2012; Rauch et al. 2012). The diagnostic yield of whole-exome sequencing (WES) varies from 30% to 50% among patients with mild to severe NDD/ID (Clark et al. 2018; Ligt et al. 2012), making WES an essential diagnostic tool for establishing a molecular diagnosis in children with NDD/ID (Clark et al. 2018; Srivastava et al. 2019). Routine reanalysis of data from undiagnosed patients often leads to a new diagnosis. The interval between the initial analysis and reanalysis typically ranges from 1 to 3 years, and among those undiagnosed patients for whom reanalysis was performed, the reported diagnostic rate ranges from 15–30%, increasing the total diagnostic yield by 10–15% (Liu et al. 2019; Costain et al. 2018). This increase in diagnostic rate can be attributed to identifying previously missed variants with advanced bioinformatics tools or collecting more informative phenotypic data from the patients; however, most of the diagnoses result from the discovery of new disease-causing genes (Liu et al. 2019; Fung et al. 2020).

In a previous single-center pilot study, we introduced a new automated, streamlined variant analysis software, referred to as EVIDENCE (Seo et al. 2020), which yielded a 42.7% diagnostic rate in patients with various clinical phenotypes. Here, we performed proband-only WES of 1065 consecutive patients with NDD/ID from 10 multimedical centers in South Korea. In this study, we implemented a daily reanalysis pipeline using EVIDENCE to enhance the diagnostic rate by quickly identifying a new gene-phenotype discovery or a change in variant pathogenicity.

## Methods

### Patients

The study enrolled an unselected series of 1065 affected individuals from 1056 nonconsanguineous families who were clinically suspected to have a genetic disorder. The patients were seen at one of the 10 clinics in South Korea from April 2018 to August 2021. Their detailed demographics, including age, sex, clinical diagnosis at the visit, family history, laboratory findings, radiologic findings, and genetic testing results, were reviewed.

Patients were included if they were strongly suspected by clinicians to have a genetic disease accompanied by neurodevelopmental delay, developmental regression, or intellectual disability. All patients were exome sequenced and analyzed as probands only after informed consent was obtained from the patients or their legal guardians after comprehensive genetic counseling. Patients and/or their guardians were counseled about the potential disclosure of medically actionable secondary findings according to the American College of Medical Genetics (ACMG) guidelines (v2.0) (Kalia et al. 2016), and they were given the option of receiving the information. The study was approved by the Institutional Review Board for Human Research of each medical center (IRB numbers: 2018-0574, 2018-0180, CHA-2018-06-008, CHH 2020-L06-01, 2020-05-040, 2020AS0186, 2020AN0332, 2020-03-031, 2020-08-003).

### Whole exome sequencing

Blood or buccal swab samples were collected from each patient, and genomic DNA was extracted from each sample. Most of the exonic regions of ~22,000 human genes were captured by one of the following 3 kits, depending on when the patient was enrolled: Agilent Sure Select kit (version C2, December 2018), Twist capture kit (Twist Bioscience HQ, San Francisco, CA, USA), or IDT xGen Exome Research Panel v2 (Integrated DNA Technologies, Coralville, Iowa, USA). Sequencing was performed using an Illumina NovaSeq6000 (San Diego, CA, USA) as 150 bp paired-end reads. The binary base call (BCL) sequence files generated by the NovaSeq6000 were converted and demultiplexed to FASTQ files. FASTQ files are aligned to the human reference genome (GRCh37/19 from NCBI, February 2009) to generate BAM files by BWA-MEM (v.0.7.17) (Li and Durbin 2009). Aligned BAM files were sorted and extracted using the statistical metric by samtools (v.1.9) (Li et al. 2009). Duplication was marked by Picard (v.2.20.8) (<http://broadinstitute.github.io/picard/>). Variant calling file were generated following the GATK best practices (GATK v.3.8) (McKenna et al. 2010). The mean depth of coverage was 125 X

(> 20 X = 97%). Detail information about sequencing quality was noted in Additional file 1: Table S1.

#### Variant analysis by EVIDENCE and daily reanalysis

Variants were annotated, filtered and prioritized using software developed in-house, EVIDENCE. EVIDENCE incorporates daily automatically updated databases, variant classification schema based on the ACMG guidelines, and a symptom similarity scoring system as previously described (Seo et al. 2020; Richards et al. 2015). Automatically updated databases include public databases, in-house variant databases and manually curated literature databases. The full list is shown in Additional file 2: Table S2. The variant classification schema incorporates the ACMG classification guidelines recommended by the ClinGen Sequence Variant Interpretation working group and assigns each variant a classification by weighing the strength of the evidence (Abou Tayoun et al. 2018; Harrison et al. 2019). The final assessment of the variant pathogenicity was determined manually by medical geneticists based on the clinical indications.

A daily reanalysis pipeline was implemented in April 2019 and patients without definite diagnosis were reanalyzed by October 2021. For patients with no clinically significant variants, only a uncertain significance of variant (VUS), potential compound heterozygous variants with one pathogenic (P)/likely pathogenic (LP) variant and one VUS, or multiple VUS reported, EVIDENCE was run on a daily basis with new annotations for the entire study period. If one or more variants were newly reclassified as P, LP or VUS with a symptom similarity score  $\geq 5$  that significantly increased the probability of being diagnosed compared to a score < 5, as described in our previous study (Seo et al. 2020), the suggested variant was rereviewed by medical geneticists and physicians. Since patient phenotypes can change over time, the pipeline updated the symptom similarity score when physicians submitted additional phenotype data after the initial analysis. Those variants excluded through the review process were not relisted until a new change was detected.

All identified variants in this study were confirmed by Sanger sequencing. A subset of the reported variants was analyzed in available family members by Sanger sequencing for segregation analysis.

#### Reporting

A positive report consisted of one or more variants that were estimated to be the disease-causing variants as follows: for an autosomal dominant disease or an X-linked disease, one heterozygous or hemizygous P/LP variant in

a known disease gene that would fit the phenotype and for an autosomal recessive disease, one homozygous P/LP variant, two P/LP compound heterozygous variants or potential compound heterozygous variants with one P/LP variant and a VUS with 2 moderate pathogenic criteria based on ACMG guidelines in a known disease gene that would fit the phenotype. An inconclusive report consisted of one heterozygous or hemizygous VUS in a known autosomal dominant or X-linked disease gene that would fit the phenotype well or potential compound heterozygous variants with one P/LP variant and one VUS with less than one moderate pathogenic criterion in a known autosomal recessive disease gene. If P/LP variants in a gene that had not been previously associated with disease defined in OMIM but reported in the literature, they were reported as inconclusive. In addition, if P/LP variants in known disease genes could fit the phenotype but needed additional phenotyping to confirm the phenotypic match, they were reported as inconclusive. Finally, if no clinically significant variant was found, a negative report was generated.

## Results

### Patient demographics

The demographic characteristics of the 1065 patients (585 males and 480 females) are shown in Table 1. The age at presentation ranged from less than age 1 month to 66 years. About eighty five percent of the patients were younger than 1 year at presentation, and 58.9% were younger than 4 months.

WES was performed at  $6.5 \pm 8.1$  years (range, 0–47 years). The average number of phenotype items according to the Human Phenotype Ontology (HPO) was  $8.6 \pm 4.6$  per patient. Abnormalities in the nervous system were most frequently observed (96.3% of patients), followed by those in the musculoskeletal system (71.9%), head and neck (57.8%), and growth abnormalities (32.1%) (Table 1).

Of the 1065 patients, 645 (60.6%) underwent separate genetic testing before WES. Furthermore, 20 patients (1.9%) underwent targeted exome sequencing, which included 4813 OMIM genes, and 286 (33.3%) underwent chromosome microarray. A total of 153 patients (14.4%) underwent single gene testing for monogenic disorders. Other genetic tests included karyotyping and/or fluorescence in situ hybridization (355 patients, 33.3%), multiplex ligation-dependent probe amplification analyses for chromosomal microdeletion or duplication syndromes (103 patients, 9.7%), and mitochondrial full genome sequencing analysis (26 patients, 2.4%). No test revealed a significant positive result.

**Table 1** Demographics of patients with neurodevelopmental delay/intellectual disability

Category	Number of patients (%)
Sex (male: female)	585 (54.9): 480 (45.1)
Onset at presentation	
Antenatal	10 (0.9%)
Neonatal (from birth to 4 months)	628 (58.9%)
Infant (from more than 4 months to 1 year)	273 (25.6%)
Early childhood (from more than 1 year to 5 years)	65 (6.1%)
Childhood (from more than 5 years to 12 years)	77 (7.2%)
Adolescent (from more than 12 years to 18 years)	10 (0.9%)
Adult	3 (0.3%)
Average number of human phenotype ontology (HPO) terms	8.6±4.6
Organ involvement	
Nervous system	1065 (100%)
Musculoskeletal and limb system	767 (71.9%)
Head or neck, including facial dysmorphism	616 (57.8%)
Growth	342 (32.1%)
Eye system	286 (26.8%)
Ear system	230 (21.6%)
Cardiovascular system	200 (18.8%)
Endocrine and metabolism/homeostasis system	185 (17.4%)
Skin	181 (17.0%)
Abnormality of prenatal development or birth	167 (15.7%)
Genitourinary system	146 (13.7%)
Gastrointestinal system	106 (9.9%)
Connective tissue system	72 (6.7%)
Blood and immune system	50 (4.7%)
Respiratory system	35 (3.3%)
Neoplasm	22 (2.1%)
Age at whole exome sequencing	6.5 year (min 0–max 47, SD 8.1)
Total number of patients with previous genetic testing	
Single gene test	645 (60.6)
Panel test	153 (14.4)
Targeted exome sequencing	13 (1.2)
Chromosome analysis or fluorescence in situ hybridization	20 (1.9)
Microarray	355 (33.3)
Multiplex ligation-dependent probe amplification	284 (26.6)
Mitochondrial full genome sequencing analysis	103 (9.7)
	26 (2.4)

Min, minimum; max, maximum; SD, standard deviation

### Positive results: molecular diagnosis by WES

#### 1. Diagnostic rate including reanalysis

The clinical and genetic information about the patients with a positive molecular diagnosis is summarized in Additional file 3: Table S3. At initial analysis, 435 variants (255 pathogenic variants, 170 likely pathogenic variants and 10 VUS) were confirmed to be causative in 401 families from 1057 unrelated

families (38.0%, 401/1056 families) throughout the study period, April 2018 to August 2021. From April 2019 to October 2021, daily updated reanalysis was performed for patients with no clinically significant variants. This reanalysis identified 49 variants in those 45 patients (Table 2). Among these variants, 34 in 31 patients were positively confirmed by reanalysis. Finally, a total of 468 variants (265 pathogenic variants, 192 likely pathogenic variants and 11 VUS) were confirmed to be causative, resulting in the molecular

**Table 2** Detailed information on 49 variants in the 45 patients identified by reanalysis

ID	Sex	Gene	Zygosity	HGVSc	Class	Molecular diagnosis	OMIM	Inheritance	Allele origin	Dx	Note
43	F	TFE3	Het	NM_006521.6:c.566C>T	P	Intellectual developmental disorder, X-linked, syndromic, with pigmentary mosaicism and coarse facies	301066	XL	Unknown	Yes	New disease
59	M	SPEN	Het	NM_015001.3:c.5806C>T	P	Radio-Tartaglia syndrome	619312	AD	Assumed de novo	Yes	New disease
62	M	ADH5	Het	NM_000671.4:c.966del	P	AMED syndrome, digenic	619151	AR	Unknown	Yes	New disease
86	F	GABRG2	Het	NM_198903.2:c.316G>A	P	Epileptic encephalopathy, early infantile, 74	618396	AD	Assumed de novo	Yes	New disease
90	F	MSL3	Het	<b>NM_078629.3:c.1226_1229del</b>	P	Basilicata-Akhtar syndrome	301032	XL	Assumed de novo	Yes	New disease
92	F	TFE3	Het	NM_006521.5:c.555A>G	LP	Intellectual developmental disorder, X-linked, syndromic, with pigmentary mosaicism and coarse facies	301066	XL	Assumed de novo	Yes	New disease
99	M	MAPK8IP3	Het	NM_001318852.1:c.1735C>T	P	Neurodevelopmental disorder with or without variable brain abnormalities	618443	AD	Assumed de novo	Yes	New disease
138	F	LTPB1	Het	<b>NM_206943.2:c.1342C&gt;T</b>	LP	Cutis laxa, autosomal recessive, type IIIE	619451	AR	Unknown	Yes	New disease
		LTPB1	Het	<b>NM_206943.2:c.4793_4794del</b>	LP				Unknown	Yes	New disease
206	M	PTPN23	Het	<b>NM_015466.3:c.345del</b>	LP	Neurodevelopmental disorder and structural brain anomalies with or without seizures and spasticity	618890	AR	Unknown	Yes	New disease
275	F	PTPN23	Het	NM_015466.3:c.4052A>G	VUS				Unknown	Yes	New disease
		PTPN23	Het	<b>NM_006521.5:c.569A&gt;G</b>	LP	Intellectual developmental disorder, X-linked, syndromic, with pigmentary mosaicism and coarse facies	301066	XL	Unknown	Yes	New disease
345	M	DIG4	Het	<b>NM_001365.4:c.1608&gt;2A&gt;G</b>	LP	Intellectual developmental disorder, 62	618793	AD	Unknown	Yes	New disease
399	F	MAP1B	Het	<b>NM_005909.4:c.6715del</b>	LP	Periventricular nodular heterotopia, 9	618918	AD	Unknown	Yes	New disease
434	M	EMC10	Het	NM_206538.3:c.343C>T	LP	Neurodevelopmental disorder with dysmorphic faces and variable seizures	619264	AR	Unknown	Yes	New disease
517	M	ZNF292	Het	<b>NM_015021.3:c.6015dup</b>	LP	Intellectual developmental disorder, autosomal dominant, 64	619264	AR	Unknown	Yes	New disease
532	M	ZMYM2	Het	<b>NM_001190964.4:c.3472C&gt;T</b>	LP	Neurodevelopmental-craniofacial syndrome with variable renal and cardiac abnormalities	619522	AD	Unknown	Yes	New disease
536	F	PHF21A	Het	<b>NM_001101802.1:c.1171A&gt;T</b>	P	Intellectual developmental disorder with behavioral abnormalities and craniofacial dysmorphism with or without seizures	618725	AD	Assumed de novo	Yes	New disease
548	F	ALKBH8	Hom	NM_001301010.1:c.1430_1438del	LP	Intellectual developmental disorder, autosomal recessive, 71	618504	AR	Trans phase	Yes	New disease
570	M	TET3	Het	<b>NM_001287491.1:c.2161_2183dup</b>	P	Beck-Fahrner syndrome	618798	AD/AR	Assumed de novo	Yes	New disease
604	M	SETD1B	Het	<b>NM_015048.1:c.5508dup</b>	LP	INTELLECTUAL DEVELOPMENTAL DISORDER WITH SEIZURES AND LANGUAGE DELAY; IDSELD	619000	AD	Unknown	Yes	New disease

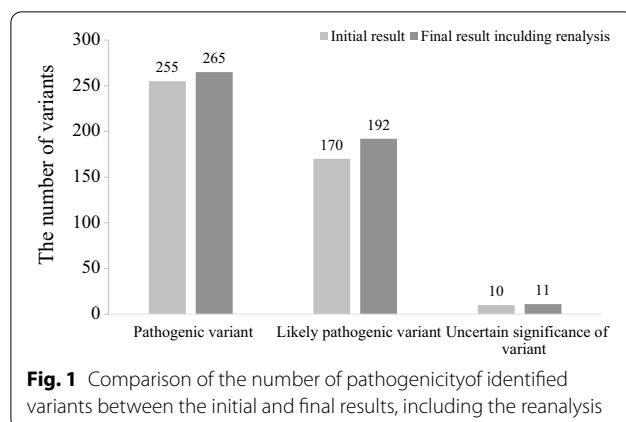
**Table 2** (continued)

ID	Sex	Gene	Zygosity	HGVSc	Class	Molecular diagnosis	OMIM	Inheritance	Allele origin	Dx	Note
656	F	SATB1	Het	NM_001195470.2:c.1574A>G <b>NM_005402.3:c.404_429dup</b>	LP	Kohlschutter-Tonz syndrome-like syndrome	619229	AD	Assumed de novo	Yes	New disease
709	F	RALA	Het	NM_001271519.1:c.2423C>T <b>NM_001014839.1:c.990dup</b>	LP	Hiatt-Neu-Cooper neurodevelopmental syndrome	619311	AD	Unknown	Yes	New disease
752	M	KCNMA1	Het	NM_001271519.1:c.2423C>T	LP	Liang-Wang syndrome	618729	AD	Unknown	Yes	New disease
776	F	NCDCN	Het	NM_001014839.1:c.990dup	LP	Neurodevelopmental disorder with infantile epileptic spasms	619373	AD/AR	Unknown	Yes	New disease
792	M	GRIA2	Het	<b>NM_001083619.1:c.1957_1958insT</b>	LP	Neurodevelopmental disorder with language impairment and behavioral abnormalities	618917	AD	Unknown	Yes	New disease
1051	F	SETD1A	Het	NM_014712.2:c.4582_2_4582_1delet	LP	Neurodevelopmental disorder with speech impairment and dysmorphic facies	619056	AD	Unknown	Yes	New disease
1053	F	ZNF292	Het	<b>NM_015021.3:c.3862dup</b>	LP	Intellectual developmental disorder, autosomal dominant 63	619188	AD	Unknown	Yes	New disease
116	M	KCNT2	Het	NM_198503.3:c.2501_2507del	LP	Developmental and epileptic encephalopathy 57	617771	AD	Unknown	Yes	Variant update
196	F	HUEW1	Het	NM_031407.6:c.12404A>C	LP	Mental retardation, X-linked syndromic, Turner type	309590	XL	Assumed de novo	Yes	Variant update
286	M	KMT2D	Het	NM_003482.3:c.10744C>T	P	Kabuki syndrome 1	147920	AD	Unknown	Yes	Variant update
379	M	HK1	Het	NM_001322365.1:c.901G>A	LP	Neurodevelopmental disorder with visual defects and brain anomalies	618547	AD	Unknown	Yes	Variant update
354	F	ACTB	Het	NM_001101.3:c.124G>A	LP	Baraitser-Winter syndrome 1	243310	AD	Assumed de novo	Yes	Clinical update
411	F	TRPM3	Het	NM_00100741.2:c.2968G>A	P	TRPM3-related intellectual disabilities with epilepsy	31278393	AD	Unknown	Not yet	Candidate gene
1052	M	TMRSS9	Het	NM_182973.2:c.1094C>A	LP	TMRSS9-related autism spectrum disorder	31943016	AR	Unknown	Not yet	Candidate gene
120	F	CHRM1	Het	NM_000738.2:c.1274T>C	LP	CHRM1-related neurodevelopmental disorder	34212451	AD	Unknown	Not yet	Candidate gene
105	F	ATAD3A	Het	NM_018188.3:c.1879C>T	VUS	Pontocerebellar hypoplasia, hypotonia, and respiratory insufficiency syndrome, neonatal lethal	618810	AR	Unknown	Not yet	New disease
39	M	ATAD3A	Het	NM_018188.3:c.1586G>A	VUS	NM_003590.5:c.383G>A	618810	AR	Unknown	Unknown	New disease
703	F	POLR2A	Het	<b>NM_000937.4:c.1832A&gt;T</b>	VUS	Neurodevelopmental disorder with or without autism or seizures	619239	AD	Unknown	Not yet	New disease
860	F	USP9X	Het	<b>NM_001039590.3:c.4796 T&gt;C</b>	VUS	Neurodevelopmental disorder with hypotonia and variable intellectual and behavioral abnormalities	618603	AD	Unknown	Not yet	Variant update
						Mental retardation, X-linked 99, syndromic, female-restricted	300968	XD	Unknown	Not yet	

**Table 2** (continued)

ID	Sex	Gene	Zygosity	HGVSc	Class	Molecular diagnosis	OMIM	Inheritance	Allele origin	Dx	Note
885	F	KDM1A	Het	NM_001009992:c.1901C>T	VUS	Cleft palate, psychomotor retardation, and distinctive facial features	616728	AD	Unknown	Not yet	Variant update
25	M	CHD7	Het	<b>NM_017780.3:c.2499-11T&gt;A</b>	VUS	CHARGE syndrome	214800	AD	Unknown	Not yet	Variant update
67	M	FGFR1	Het	<b>NM_001174067.1:c.1099G&gt;A</b>	VUS	Osteoglophonic dysplasia	166250	AD	Unknown	Not yet	Variant update
514	F	EBF3	Het	<b>NM_001005463.3:c.373G&gt;A</b>	VUS	Hypotonia, ataxia, and delayed development syndrome	617330	AD	Unknown	Not yet	Variant update
902	F	SPTBN1	Het	<b>NM_003128.2:c.5708G&gt;A</b>	VUS	Developmental delay, impaired speech, and behavioral abnormalities	619475	AD	Unknown	Not yet	Variant update
979	F	SPTBN1	Het	<b>NM_003128.3:c.2471T&gt;C</b>	VUS	Developmental delay, impaired speech, and behavioral abnormalities	619475	AD	Unknown	Not yet	Variant update
990	M	NSF	Het	<b>NM_006178.4:c.1359G&gt;A</b>	VUS	Developmental and epileptic encephalopathy 96	619340	AD	Unknown	Not yet	Variant update

P, pathogenic; LP, Likley pathogenic; VUS, uncertain significance of variant; AD, autosomal dominant; AR, autosomal recessive; XL, X-linked; Het, heterozygous; Hom, Dx, diagnosis; Bold: novel variant



**Fig. 1** Comparison of the number of pathogenicity of identified variants between the initial and final results, including the reanalysis

diagnosis of 433 families from 1056 unrelated families (41%, 433/1056 families) (Figs. 1 and 2).

## 2. Characteristics of variant type and disease

The most common variant type was missense (45.1%), followed by frameshift (23.2%), stop gain (18.4%), canonical splice site (8.5%), and other (4.8%) variants. Remarkably, 47.5% (220/463) of all confirmed variants were not previously described in any public database.

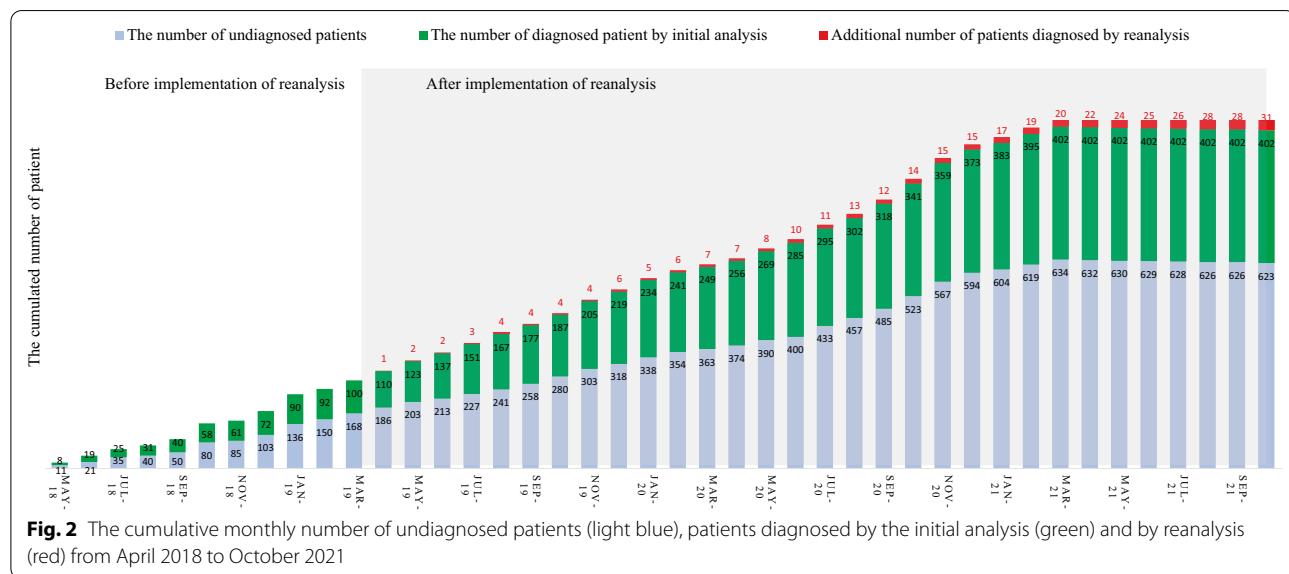
Among 433 unrelated patients with a molecular diagnosis, the parents of 279 underwent Sanger sequencing to confirm the segregation pattern of the identified variant. Two hundred three variants were confirmed to be assumed de novo (i.e., biological relationship unconfirmed). Eighty variants from 44 patients were observed in a *trans* pattern with the other variant. Six variants were inherited from

symptomatic mothers or fathers. Nine variants of the X chromosome were inherited from asymptomatic mothers. Thirteen variants inherited from an asymptomatic parent were attributed to incomplete penetrance. Two variants showed maternal mosaicism. The number of patients with variants and the inheritance of these variants are summarized in Fig. 3.

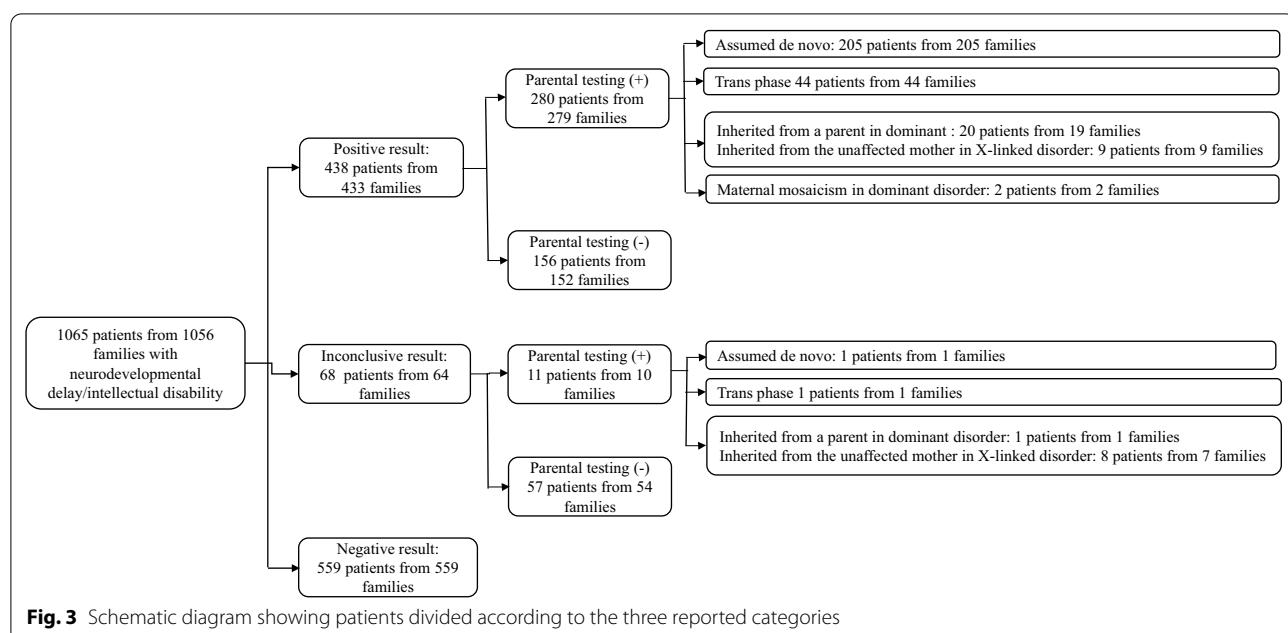
A total of 279 Mendelian disease genes were found in the 433 families, including autosomal dominant ( $N=196$ , 70.3%), autosomal recessive ( $N=53$ , 19.0%), X-linked ( $N=24$ , 8.6%), and autosomal dominant and/or recessive ( $N=6$ , 2.1%) inheritance patterns. Most genes were reported once (202 genes), twice (48 genes) or three times (11 genes). *PTPN11* (Noonan syndrome 1, OMIM 163950) and *ARID1B* (Coffin-Siris syndrome 1, OMIM 135900) were recurrently reported in 11 patients and 10 patients, respectively. Subsequently, *KMT2D* (9 patients), *NF1* (9 patients), *CTNNB1* (7 patients), *ZEB2* (6 patients), *PRRT2* (6 patients), *ANKRD11* (5 patients), *BRAF* (5 patients), *SCN1A* (5 patients), *FOXP1* (5 patients), *ASXL3* (5 patients), *MECP2* (4 patients), *PIK3CA* (4 patients), *EP300* (4 patients), *TUBB3* (4 patients), *WDR45* (4 patients), and *MED13L* (4 patients) were recurrently reported.

## 3. Dual diagnosis and confirmation of VUSs

Two patients among the 433 unrelated patients with a positive result received a dual molecular genetic diagnosis. A female patient (ID: 196) showed global developmental delay (GDD), dysmorphic facial features, microphthalmia, micrognathia, leukodystrophy, microcephaly, hypotonia, spasticity, and club foot. She was diagnosed with peroxisome biogen-



**Fig. 2** The cumulative monthly number of undiagnosed patients (light blue), patients diagnosed by the initial analysis (green) and by reanalysis (red) from April 2018 to October 2021



**Fig. 3** Schematic diagram showing patients divided according to the three reported categories

esis disorder 2A (Zellweger) and Mental retardation, X-linked syndromic, Turner type caused by compound heterozygous variants (NM\_001300789.1: c.240A>C and c.1622A>G) in the *PEX5* gene and an assumed de novo variant (NM\_031407.6: c.12404A>C) in the *HUEW1* gene. A female patient (ID: 98) presenting with hydrocephalus, autistic features, GDD, microcephaly, hypotonia, and dysmorphic facial features was diagnosed with congenital contractures of the limbs and face, hypotonia, and developmental delay, and neurodevelopmental disorder with severe motor impairment and absent language caused by an assumed de novo variant (NM\_052867.2:c.3731 T>G) in the *NALCN* gene and a known pathogenic variant (NM\_138615.2: c.2344C>T) in the *DHX30* gene, respectively. These genetic variants were associated with either nonoverlapping clinical presentations or contributed to one major phenotype.

Among the 72 VUSs identified in 68 patients that were determined to involve clinically relevant genes based on OMIM data or recent PubMed publications, positive results were obtained for 11. Six of these 11 VUSs were confirmed to be in trans phase with other LP or P variants in genes associated with autosomal recessive disorder and were highly matched with disease-relevant symptoms. Three VUSs (ID: 206, 210, and 697) with other LP or P variants also showed highly specific phenotypes associated with each disease; however, phase status was undetermined due to a lack of familial segregation

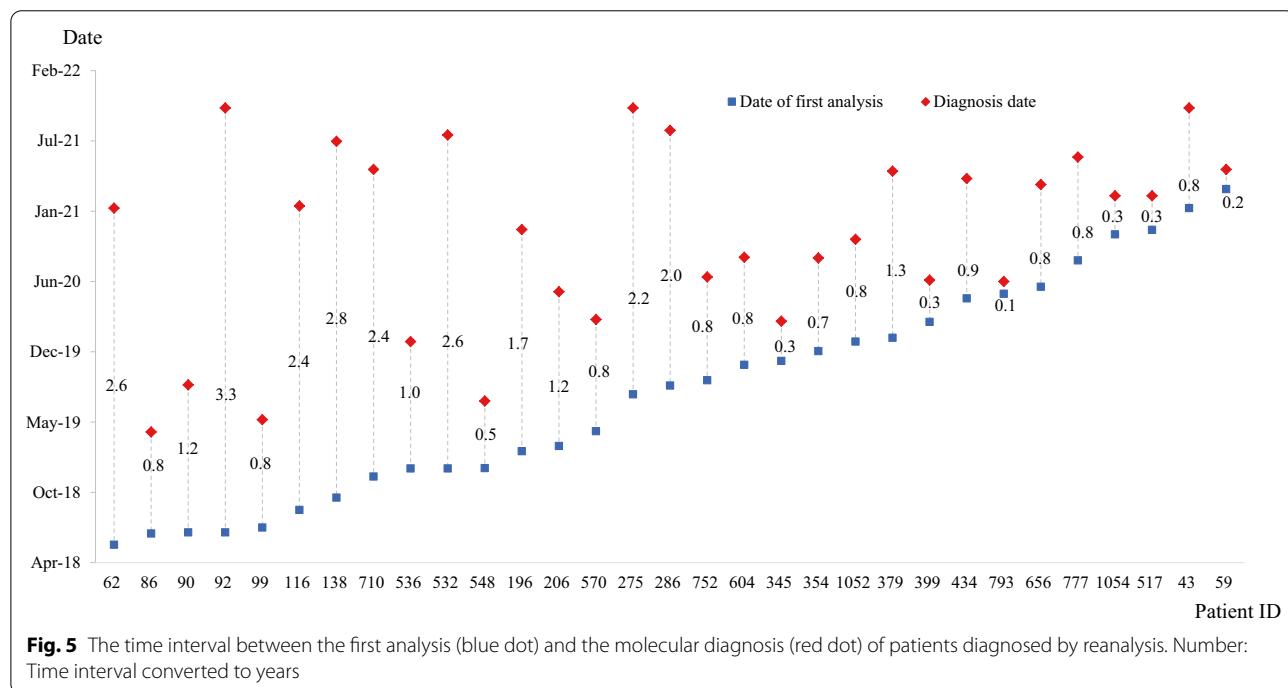
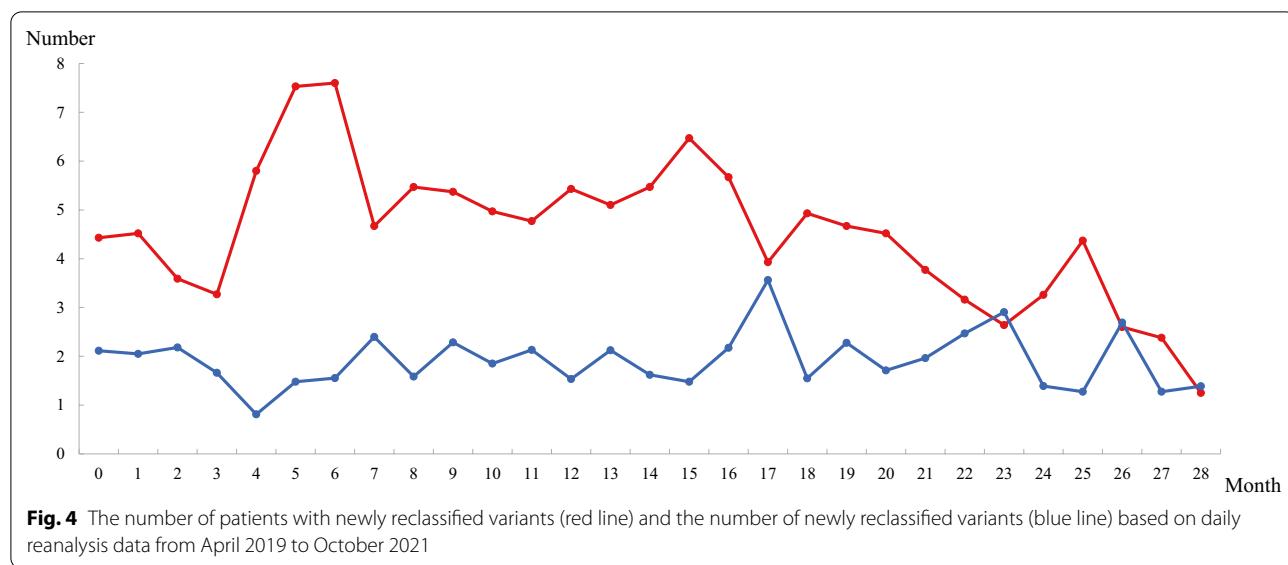
analysis. Importantly, the VUS of 2 patients was confirmed by clinicians using other genetic testing methods and by assessing treatment responses. A male patient (ID: 143) had maternal UPD 9 encompassing a region containing the VUS (Seo et al. 2020). A male patient (ID: 71) with a homozygous variant in the *PREPL* gene responded to medical treatment after molecular diagnosis (Kim et al. 2020).

#### Inconclusive results

Sixty-four variants, including 1 P, 2 LP and 61 heterozygous VUSs, were assessed as inconclusive, i.e., insufficient to be responsible for the patients' phenotypes. Fifty-five variants had not been previously reported. These variants were identified in 61 patients from 57 unrelated families (Additional file 4: Table S4) and associated with 36 autosomal dominant, 6 autosomal recessive and 12 X-linked disorders. Furthermore, we identified 1 assumed de novo LP variant and 7 heterozygous variants in 7 genes that have not been registered as OMIM morbid genes but have been described in a few studies (Additional file 4: Table S4).

#### Summary of reanalyzed results and description of patients diagnosed with a new disease

From April 2019 to October 2021, daily updated reanalysis reported an average of  $2.1 \pm 0.7$  newly reclassified variants in an average of  $4.1 \pm 1.7$  patients per day among all undiagnosed patients (Fig. 4). The time interval between the first analysis and the molecular diagnosis by reanalysis was  $1.2 \pm 0.9$  years (from a minimum of 1 month to



a maximum of 3.3 years, Fig. 5). This time interval was shorter among the patients enrolled during the latter part of the study period. Table 2 shows 49 variants identified in the 45 patients by reanalysis (Table 2). Diagnosis by reanalysis was possible in 26 patients due to the discovery of a new gene-disease relationship. They were diagnosed with 23 different diseases that were discovered after 2019. In addition, the updated variant reclassification and updated phenotype description by physicians helped

diagnose 4 patients and 1 patient, respectively. Molecular diagnoses by reanalysis were achieved for 31 patients. In addition, the results were inconclusive in the remaining 14 patients (1.3%, 14/1056) with 2 P/LP variants in 2 candidate genes and 12 VUSs due to lack of additional in vivo or in vitro evidence.

In 6 patients, the diagnosis was obtained by automated daily updated analysis approximately 100 days after the initial report, e.g., in a male patient (ID: 59) who

presented with hypotonia, facial dysmorphism including thin ear helix, prominent ear, prominent upper lip vermillion, short philtrum, bulbous nose, flat nasal root, hypertelorism and thick eyebrow, hearing loss, optic atrophy, nephrocalcinosis, autistic features, and global developmental delay. The patient's initial March 2021 report was negative; however, a *SPEN* variant was first reported in March 2021 in 13 individuals with similar phenotypes (Radio et al. 2021). On May 3, 2021, this variant was registered as a new entry associated with Radio-Tartaglia syndrome (OMIM 619312). The patient was found to have the P variant in *SPEN* (NM\_015001.3: c.5806C>T). Therefore, 2 months after the initial report, he was diagnosed with Radio-Tartaglia syndrome (OMIM 619312) by reanalysis. A male patient (ID: 345) presented with intellectual disability, autistic features, periventricular leukomalacia, and facial dysmorphism, including prominent ears, flat nasal root, prominent upper vermillion, and hypertelorism. The initial report was negative in Nov 2019. On Feb 26, 2020, intellectual developmental disorder 62 (OMIM 618793) caused by the *DLG4* variant was registered as a new entry. This patient was found to have the LP variant in *DLG4* (NM\_001365.4: c.1608-2A>G). Therefore, 3 months after the initial report, he was diagnosed with intellectual developmental disorder 62 (OMIM 618793) by reanalysis. A female patient (ID: 399) presented with GDD, heterotopia, failure to thrive, seizures, and facial dysmorphism. She had the LP variant in *MAP1B* (NM\_005909.4: c.6715del), but the initial report was inconclusive in Feb 2020 since this gene-phenotype association had been reported but not yet registered in

OMIM (Julca et al. 2019). On June 22, 2020, periventricular nodular heterotopia 9 (OMIM 618918) caused by the *MAP1B* variant was registered as a new entry. Therefore, 3 months after the initial report, this patient was diagnosed with periventricular nodular heterotopia 9 (OMIM 618918) by reanalysis. A male patient (ID: 793) presented with GDD, strabismus, and facial dysmorphism. He had the LP variant in *GRIA2* (NM\_000826.6: c.1958\_1960delCCAinsTCTACAGCAC), but the initial report was inconclusive in May 2020 since this gene-phenotype association was reported in July 2019 but had not been registered in OMIM (Salpietro et al. 2019). On June 18, 2020, neurodevelopmental disorder with language impairment and behavioral abnormalities (OMIM 618917) caused by the *GRIA2* variant was registered as a new entry. Therefore, 1 month after the initial report, the patient was diagnosed with neurodevelopmental disorder with language impairment and behavioral abnormalities (OMIM 618917) by reanalysis. Finally, one male (ID: 517) and one female (ID: 1054) patient presenting with GDD received negative initial reports in December 2020. The *ZNF292* variant was first reported in 28 families with intellectual disabilities in March 2020 (Mirzaa et al. 2020). On February 2021, this variant was registered as a new entry associated with intellectual developmental disorder, autosomal dominant 64 (OMIM 619188). The two patients had LP variants in *ZNF292* (NM\_015021.3: c.6015dup and NM\_015021.3: c.3862dup). Therefore, 2 months after the initial report, they were diagnosed with intellectual developmental disorder, autosomal dominant 64 (OMIM 619188), by reanalysis.

**Table 3** The 16 patients in whom clinical management was changed after the genetic diagnosis

Patient ID	Gene	Disease	Medical treatment
35	<i>SCN8A</i>	Epileptic encephalopathy, early infantile, 13	Na channel blocker including oxcarbazepine and phenytoin
71	<i>PREPL</i>	Myasthenic Syndrome, Congenital, 22	Pyridostigmine
159	<i>SLC2A1</i>	GLUT1 Deficiency Syndrome 1	Ketogenic diet
228	<i>SCN8A</i>	Epileptic encephalopathy, early infantile, 13	Na channel blocker including oxcarbazepine and phenytoin
269	<i>KCNQ2</i>	Epileptic encephalopathy, early infantile, 7	Na channel blocker including oxcarbazepine and phenytoin
278	<i>CPS1</i>	Carbamoylphosphate synthetase I deficiency	Low protein diet
285	<i>SCN1A</i>	Epileptic encephalopathy, early infantile, 6	Valprotic acid, topiramate
314	<i>SCN2A</i>	Epileptic encephalopathy, early infantile, 11	Valprotic acid, topiramate
363	<i>TH</i>	Segawa syndrome, recessive	Levodopa
402	<i>ASS1</i>	Citrullinemia	Low protein diet
442	<i>KCNQ2</i>	Epileptic encephalopathy, early infantile, 7	Na channel blocker including oxcarbazepine and phenytoin
518	<i>PRRT2</i>	Seizures, benign familial infantile, 2	Oxcarbazepine
519	<i>SCN1A</i>	Febrile seizures, familial, 3A	Valprotic acid, topiramate
229	<i>NF1</i>	Neurofibromatosis Type 1	Selumetinib
257	<i>TOR1A</i>	Dystonia-1, torsion	Deep brain stimulation
288	<i>ARSA</i>	Metachromatic leukodystrophy	Bone marrow transplantation candidate
383	<i>NF1</i>	Neurofibromatosis Type 1	Selumetinib

### Clinical actionability after diagnosis

The genetic diagnosis changed the clinical management in 4.0% of the diagnosed patients (16/402 pts) (Table 3). After diagnosis, eight patients with seizure disorder (ID: 35, 228, 269, 285, 314, 442, 518, and 519) changed antiepileptic drug. Ketogenic diet was started in a patient 159 with GLUT1 deficiency syndrome 1. Patient 71 and 363 with myasthenic syndrome, congenital, 22 and segawa syndrome, recessive received pyridostigmine and levodopa beginning at age 2.6 years and 10 months, respectively (Kim et al. 2020). Patient 278 and 402 with urea cycle disorder were managed with low protein diet. Patient 229 and 383 with neurofibromatosis type 1 having inoperable plexiform neurofibroma was treated with selumetinib as in a clinical trial (<https://cris.nih.go.kr/cris/search/detailSearch.do/19080>). In case of patient 257, dystonia was refractory to medical treatment, and deep brain stimulation surgery was performed.

### Secondary findings

Forty-two variants, including 16 P and 19 LP variants, were identified in the 17 genes recommended to be reported as secondary findings by the ACMG guidelines (v2.0) in 35 of the 1056 patients (3.3%). Twenty-nine variants (16 P and 13 LP variants) had been previously reported to be P or LP. Pathogenic variants in the *BRCA2* gene were the most commonly identified, but recurrent variants were not detected. Genetic counseling was provided for the patients and their family members, and appropriate surveillance was conducted depending on the identified genes.

### Discussion

WES is an effective diagnostic approach with a higher diagnostic rate than gene panels or CMA (Clark et al. 2018; Srivastava et al. 2019). Studies employing WES for investigating neurodevelopmental disorders have revealed that the diagnostic yield varies from 30 to 50% (Lecoquierre et al. 2019; Clark et al. 2018; Ligt et al. 2012; Kim et al. 2019). The diagnostic yield of ~40% found here indicates that WES is a valuable tool for diagnosing patients with NDD/ID and is comparable to that observed in other studies.

Most of our patients had syndromic features with an average of 8 organ system abnormalities, indicating widely variable phenotypic heterogeneity. In line with these clinical observations, molecular diagnosis revealed a wide range of genetic diseases associated with P/LP variants in 279 different Mendelian disorder genes. Most genes were reported once (202 genes) or twice (48 genes), and only a small portion of genes (30 genes) were recurrently reported in our study.

In previous retrospective studies, reanalysis has increased the diagnostic rate of 10% over a period of

18–36 months, with total diagnostic rate as ~40% (Fung et al. 2020). In our study, the diagnostic rate was initially 38.0%, and reanalysis increased the diagnostic rate by 3%, which was lower than the increase observed in previous studies (Liu et al. 2019; Costain et al. 2018; Fung et al. 2020; Wenger et al. 2017). The main reason for this lower enhancement is that our automated system was updated daily and applied the new information to the analysis immediately after each new gene-phenotype causality was reported. As in the prospective study, this daily update allowed diagnoses based on the initial analyses of patients enrolled in the latter part of the study. Indeed, the interval between the date of ordering for WES and the date of the gene-disease causality was originally reported was less than 1 year in the 17 identified diseases. As shown in Fig. 2, the relatively consistent diagnostic rate enhancement by reanalysis reflects the effect of daily updates of the analysis system. In addition, this time interval was shorter among the patients enrolled during the latter part of the study period, as shown in Fig. 5, demonstrating the excellent performance of this automated, daily reanalysis approach.

It has recently been recommended that reanalysis of an individual's genome or exome sequencing data should be performed every 1–2 years until diagnosis, or more frequently if their phenotype evolves, considering the cost and resources used (Costain et al. 2018; Ewans et al. 2018). However, it is not easy or convenient to regularly check up on and reanalyze every undiagnosed patient, and updated diagnoses can be easily missed. Our automated daily updated analysis system can effectively reanalyze data from undiagnosed patients, minimizing missed diagnoses.

A variety of reasons for the improved diagnostic yield obtained by reanalysis have been reported, including new gene-disease associations and literature updates, updated and clarified patient phenotypes, additional sequencing for trios/other affected individuals and family members, improvements in sequencing data by resequencing strategies, upgraded bioinformatics tools, and research collaborations (Liu et al. 2019; Fung et al. 2020; Won et al. 2020). The main reason for obtaining additional diagnoses by reanalysis has been attributed to the discovery of new gene-disease associations (Liu et al. 2019; Fung et al. 2020). Approximately 250 novel gene-disease and 9200 novel variant-disease associations are reported yearly (Wenger et al. 2017). New entries are uploaded daily to OMIM or ClinVar, which is one of the most important databases of diseases and variants. In the current study, most of the reanalyzed patients were diagnosed due to newly discovered gene-disease associations. Six patients were diagnosed with a new disease within approximately 100 days of the first analysis. In addition, among two patients with dual molecular genetic

diagnoses, one patient (ID: 196) was initially diagnosed with peroxisome biogenesis disorder 2A. However, the patient's data were reanalyzed due to compound heterozygous variants including one LP variant and one VUS of the *PEX5* gene. One VUS of the *HUEW1* gene with high similarity scores were selected, leading to a dual genetic diagnosis as it was confirmed to be assumed de novo. Therefore, the reanalysis system can be effective, identifying a new diagnosis in a patient who was previously clinically diagnosed but with molecularly inconclusive results.

Reanalysis is difficult in conventional clinical laboratories because it is time-consuming and labor-intensive. In addition, frequent changes in analytical pipelines may hinder routine work. In our reanalysis system, an average of 2.1 variants per patient were newly reclassified in an average of 4 undiagnosed patients daily up to 800 days after the initial analysis. Since variants that were newly uploaded and reviewed the previous day were excluded from the subsequent analysis, the number of newly reclassified variants tended to decrease after 800 days. Therefore, an automated pipeline-based daily updated system significantly reduced the workload required for the periodic reanalysis of undiagnosed patients.

Several studies comparing the diagnostic yield of proband-only WES with trio WES in patients with a complex phenotype have shown that trio WES provides an incremental gain of 10–15%, mainly because de novo variants can be easily identified (Srivastava et al. 2019; Tan et al. 2019). Moreover, the better efficiency of trio WES versus proband-only WES was attributed to the nearly twofold reduction in the number of variants selected for curation due to the use of parental information (Tan et al. 2019). Although candidate variants identified by proband-only WES require additional targeted Sanger sequencing to confirm segregation, information that is intrinsically obtained from trio WES, the overall diagnostic yield of 41% in this study using proband-only WES was in line with previous studies of trio WES for patients with NDD/ID (Clark et al. 2018; Srivastava et al. 2019). A trio-based approach has been more effective in discovering novel gene-disease associations and in increasing confidence in the role of variants of unknown significance (Bertoli-Avella et al. 2021; Farwell et al. 2015). However, in a clinical setting, trio testing would not be feasible due to the increased costs of parental WES and the possibility that parental DNA is not available. In addition, the curating times and costs of WES analysis were reduced due to automatic variant interpretation by EVIDENCE (Seo et al. 2020; Kim et al. 2021), thus suggesting that proband-only WES by EVIDENCE is a comparable diagnostic approach for these patients in a routine clinical setting.

This study had several limitations. First, family member testing could not be performed in 279 of the 433 patients diagnosed with genetic diseases due to unavailable samples. Second, maternity and paternity confirmation was not performed due to Confucian ideals of South Korea. Most de novo occurrences were assumed. Third, the classification of variants may change as information is updated in the daily-updated databases. Therefore, as criteria for each variant can be added or removed based on updated information, the variant classification change. Fourth, not all patients with inconclusive or negative results were reanalyzed, and only those with variants newly reclassified by automatic reanalysis were rereviewed. However, there may be differences in variant classification depending on how the cutoff criteria for selecting a variant are set based on the ACMG guidelines and additional databases. In addition, symptom similarity may be underestimated if information regarding disease or patient symptoms is limited. Finally, since a manual review was not performed for the variants not selected by the automatic reanalysis system, it is possible that the diagnosis was missed in some patients.

## Conclusion

In conclusion, this study was the first to consider the diagnostic and clinical utility of WES in a large, multi-medical center cohort group of patients with NDD/ID in South Korea. Our results suggest that proband-only WES could be useful for the genetic diagnosis of patients with NDD/ID with diverse genetic causes. Daily updated databases and reanalysis systems can provide a higher diagnostic rate using initial analysis and contribute to the rapid diagnosis of undiagnosed patients by reflecting the latest medical knowledge.

## Abbreviations

WES: Whole-exome sequencing; NDD: Neurodevelopmental delay; ID: Intellectual disability; GDD: Global developmental delay; CMA: Chromosomal microarray; ACMG: American College of Medical Genetics; VUS: Uncertain significance of variant; P: Pathogenic; LP: Likely pathogenic; HPO: Human Phenotype Ontology.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s10020-022-00464-x>.

**Additional file 1:** Table S1. Sequencing quality information for 1065 patient.

**Additional file 2:** Table S2. Database used for variant annotation and interpretation.

**Additional file 3:** Table S3. Detailed information on patients with a positive result.

**Additional file 4:** Table S4. Detailed information on patients with inconclusive result.

## Acknowledgements

We deeply appreciate the patients and their families for participating in this study.

## Authors' contributions

GHS drafted and revised the manuscript critically for important intellectual content and substantially contributed to the interpretation of data. HL, JL, HH, and YKC substantially contributed to the development of EVIDENCE and automatic reanalysis systems. MK, YC, JC, IC, SR, KYC, YK, CKC, SJK, JL, EK, JHB, HJY, YS, AO, WJK-treated patients performed all clinical analyses. MSY, BHL and BLE designed the study, revised the manuscript critically for important intellectual content, and approved the final version to be published. All authors read and approved the final manuscript.

## Funding

Institute for Information and Communications Technology Promotion (IITP) grant funded by the Korean government (MSIT) (2018-0-00861, Intelligent SW Technology Development for Medical Data Analysis).

## Availability of data and materials

All variant and phenotype data supporting findings of this study are available within the manuscript and supplement tables. The pathogenic variants identified in this study have been submitted to ClinVar with Accession Numbers (SCV002011902-SCV002012361, SCV002011867, SCV002011868, SCV002011869, and SCV002011870). The raw data of whole-exome sequencing of the patient in this study are not publicly available to protect participant confidentiality, but they are available from the corresponding author on reasonable request.

## Declarations

### Ethical approval and consent to participate

The study was approved by the Institutional Review Board for Human Research of each medical center (IRB numbers: 2018-0574, 2018-0180, CHA-2018-06-008, CHH 2020-L06-01, 2020-05-040, 2020AS0186, 2020AN0332, 2020-03-031, 2020-08-003).

### Consent for publication

Not applicable.

### Competing interests

The authors declare no conflicts of interest.

### Author details

<sup>1</sup>Billion Inc., Seoul, South Korea. <sup>2</sup>Department of Pediatrics, Asan Medical Center Children's Hospital, University of Ulsan College of Medicine, 88, Olympic-ro 43-Gil, Songpa-Gu, Seoul 05505, South Korea. <sup>3</sup>Medical Genetics Center, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea. <sup>4</sup>Department of Pediatrics, CHA Bundang Medical Center, CHA University School of Medicine, Seongnam, South Korea. <sup>5</sup>Department of Pediatrics, Chungnam National University Sejong Hospital, Sejong, South Korea. <sup>6</sup>Department of Pediatrics, Pusan National University Children's Hospital, Yangsan, Korea. <sup>7</sup>Department of Pediatrics, Inha University Hospital, Inha University College of Medicine, Incheon, South Korea. <sup>8</sup>Department of Pediatrics, Korea University Guro Hospital, 148 Gurodong-ro, Guro-gu, Seoul 08308, South Korea. <sup>9</sup>Department of Pediatrics, Hallym University College of Medicine, Hwasung, South Korea. <sup>10</sup>Department of Pediatrics, Soonchunhyang University Bucheon Hospital, Soonchunhyang University School of Medicine, Bucheon, South Korea. <sup>11</sup>Department of Pediatrics, Chungbuk National University Hospital, Chungbuk National University College of Medicine, Cheongju, South Korea. <sup>12</sup>Department of Laboratory Medicine, EONE Laboratories, Seoul, South Korea.

Received: 23 December 2021 Accepted: 15 March 2022

Published online: 26 March 2022

## References

- Abou Tayoun AN, Pesaran T, DiStefano MT, Oza A, Rehm HL, Biesecker LG, Harrison SM. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Hum Mutat*. 2018;39(11):1517–24.
- Bertoli-Avella AM, Kandaswamy KK, Khan S, Ordonez-Herrera N, Tripolszki K, Beetz C, Rocha ME, Urzi A, Hotakainen R, Leubauer A, et al. Combining exome/genome sequencing with data repository analysis reveals novel gene-disease associations for a wide range of genetic disorders. *Genet Med*. 2021;23(8):1551–68.
- Casanova EL, Gerstner Z, Sharp JL, Casanova MF, Feltus FA. Widespread genotype-phenotype correlations in intellectual disability. *Front Psychiatry*. 2018;9:535.
- Clark MM, Stark Z, Farnaes L, Tan TY, White SM, Dimmock D, Kingsmore SF. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genom Med*. 2018;3:16.
- Costain G, Jobling R, Walker S, Reuter MS, Snell M, Bowdin S, Cohn RD, Dupuis L, Hewson S, Mercimek-Andrews S, et al. Periodic reanalysis of whole-genome sequencing data enhances the diagnostic advantage over standard clinical genetic testing. *Eur J Hum Genet*. 2018;26(5):740–4.
- de Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, Kroes T, Vulto-van Silfhout AT, Koolen DA, de Vries P, Gilissen C, et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med*. 2012;367(20):1921–9.
- Ewans LJ, Schofield D, Shrestha R, Zhu Y, Gayevskiy V, Ying K, Walsh C, Lee E, Kirk EP, Colley A, et al. Whole-exome sequencing reanalysis at 12 months boosts diagnosis and is cost-effective when applied early in Mendelian disorders. *Genet Med*. 2018;20(12):1564–74.
- Farwell KD, Shahmirzadi L, El-Khechen D, Powis Z, Chao EC, Tippin Davis B, Baxter RM, Zeng W, Mroske C, Parra MC, et al. Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. *Genet Med*. 2015;17(7):578–86.
- Fung JLF, Yu MHC, Huang S, Chung CCY, Chan MCY, Pajusalu S, Mak CCY, Hui VCC, Tsang MHY, Yeung KS, et al. A three-year follow-up study evaluating clinical utility of exome sequencing and diagnostic potential of reanalysis. *NPJ Genom Med*. 2020;5:37.
- Harrison SM, Biesecker LG, Rehm HL. Overview of Specifications to the ACMG/AMP Variant Interpretation Guidelines. *Curr Protoc Hum Genet*. 2019;103(1):e93.
- Julca DM, Diaz J, Berger S, Leon E. MAP1B related syndrome: case presentation and review of literature. *Am J Med Genet A*. 2019;179(9):1703–8.
- Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, Herman GE, Hufnagel SB, Klein TE, Korf BR, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med*. 2017;19(2):249–55.
- Kim SH, Kim B, Lee JS, Kim HD, Choi JR, Lee ST, Kang HC. Proband-only clinical exome sequencing for neurodevelopmental disabilities. *Pediatr Neurol*. 2019;99:47–54.
- Kim SY, Lee S, Seo GH, Kim BJ, Oh DY, Han JH, Park MK, Lee SM, Kim B, Yi N, et al. Powerful use of automated prioritization of candidate variants in genetic hearing loss with extreme etiologic heterogeneity. *Sci Rep*. 2021;11(1):19476.
- Kim MJ, Yum MS, Seo GH, Lee Y, Jang HN, Ko TS, Lee BH. Clinical application of whole exome sequencing to identify rare but remediable neurologic disorders. *J Clin Med*. 2020;9(11):3724.
- Lecoquierre F, Duffourd Y, Vitobello A, Bruel AL, Urteaga B, Coubes C, Garret P, Nambot S, Chevarin M, Jouan T, et al. Variant recurrence in neurodevelopmental disorders: the use of publicly available genomic data identifies clinically relevant pathogenic missense variants. *Genet Med*. 2019;21(11):2504–11.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754–60.

- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9.
- Liu P, Meng L, Normand EA, Xia F, Song X, Ghazi A, Rosenfeld J, Magoulas PL, Braxton A, Ward P, et al. Reanalysis of clinical exome sequencing data. *N Engl J Med*. 2019;380(25):2478–80.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Gariumella K, Altshuler D, Gabriel S, Daly M, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297–303.
- Mirzaa GM, Chong JX, Piton A, Popp B, Foss K, Guo H, Harrapaul R, Xia K, Scheck J, Aldinger KA, et al. De novo and inherited variants in ZNF292 underlie a neurodevelopmental disorder with features of autism spectrum disorder. *Genet Med*. 2020;22(3):538–46.
- Moeschler JB, Shevell M. Comprehensive evaluation of the child with intellectual disability or global developmental delays. *Pediatrics*. 2014;134(3):e903–918.
- Radio FC, Pang K, Ciolfi A, Levy MA, Hernández-García A, Pedace L, Pantaleoni F, Liu Z, de Boer E, Jackson A, et al. SPEN haploinsufficiency causes a neurodevelopmental disorder overlapping proximal 1p36 deletion syndrome with an episignature of X chromosomes in females. *Am J Hum Genet*. 2021;108(3):502–16.
- Rauch A, Wieczorek D, Graf E, Wieland T, Endele S, Schwarzmayr T, Albrecht B, Bartholdi D, Beygo J, Di Donato N, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet*. 2012;380(9854):1674–82.
- Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405–24.
- Ropers HH. Genetics of early onset cognitive impairment. *Annu Rev Genomics Hum Genet*. 2010;11:161–87.
- Salpeter V, Dixon CL, Guo H, Bello OD, Vandrovčová J, Efthymiou S, Maroofian R, Heimer G, Burglen L, Valence S, et al. AMPA receptor GluA2 subunit defects are a cause of neurodevelopmental disorders. *Nat Commun*. 2019;10(1):3094.
- Seo GH, Kim T, Choi IH, Park JY, Lee J, Kim S, Won DG, Oh A, Lee Y, Choi J, et al. Diagnostic yield and clinical utility of whole exome sequencing using an automated variant prioritization system, EVIDENCE. *Clin Genet*. 2020;98(6):562–70.
- Srivastava S, Love-Nichols JA, Dies KA, Ledbetter DH, Martin CL, Chung WK, Firth HV, Frazier T, Hansen RL, Prock L, et al. Meta-analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. *Genet Med*. 2019;21(11):2413–21.
- Srouf M, Shevell M. Genetics and the investigation of developmental delay/intellectual disability. *Arch Dis Child*. 2014;99(4):386–9.
- Study TD. Prevalence and architecture of de novo mutations in developmental disorders. *Nature*. 2017;542(7642):433–8.
- Tan TY, Lunke S, Chong B, Phelan D, Fanjul-Fernandez M, Marum JE, Kumar VS, Stark Z, Yeung A, Brown NJ, et al. A head-to-head evaluation of the diagnostic efficacy and costs of trio versus singleton exome sequencing analysis. *Eur J Hum Genet*. 2019;27(12):1791–9.
- Vasudevan P, Suri M. A clinical approach to developmental delay and intellectual disability. *Clin Med (Lond)*. 2017;17(6):558–61.
- Vissers LE, Gilissen C, Veltman JA. Genetic studies in intellectual disability and related disorders. *Nat Rev Genet*. 2016;17(1):9–18.
- Wenger AM, Guturu H, Bernstein JA, Bejerano G. Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers. *Genet Med*. 2017;19(2):209–14.
- Won D, Kim SH, Kim B, Lee ST, Kang HC, Choi JR. Reanalysis of genomic sequencing results in a clinical laboratory: advantages and limitations. *Front Neurol*. 2020;11:612.
- Zablotsky B, Black LI, Maenner MJ, Schieve LA, Danielson ML, Bitsko RH, Blumberg SJ, Kogan MD, Boyle CA. Prevalence and trends of developmental disabilities among children in the United States: 2009–2017. *Pediatrics*. 2019;144(4):e20190811.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](http://biomedcentral.com/submissions)

